ABSTRACT
Nguyen and Yamada [NY’13] proposed an adaptive algorithm for fast and stable extraction of the first generalized Hermitian eigenvector and mentioned the extension to the first r generalized eigenvector extraction based on the nested orthogonal complement structure [NTY’12]. However, we recently found that the estimates of the eigenvectors are not expressed ideally in the time-varying coordinate system and can change drastically in a certain situation, which may cause numerical instability. In this paper, we propose a new expression of the estimates along with time-varying coordinate system. This modification can be done efficiently with additional multiplications of orthogonal complement matrices. Numerical experiments show that the modified scheme has better stability compared with the original scheme [NTY’12].

Index Terms— Generalized Hermitian eigenvalue problem (GHEP), Adaptive algorithm, Stabilization, Orthogonal complement matrix, Nested orthogonal complement structure

1. INTRODUCTION
Generalized Hermitian eigenvalue problem (GHEP) is an estimation problem, for a pair of Hermitian positive definite matrices \((R_y, R_z) \in \mathbb{C}^{N \times N} \times \mathbb{C}^{N \times N}\), of vectors \(v_i \in \mathbb{C}^N \setminus \{0\}\) satisfying
\[
R_y v_i = \lambda_i R_z v_i \quad \text{s.t.} \quad v_i^H R_z v_j = \delta_{i,j} (i, j = 1, 2, \ldots, N),
\]
where \(\lambda_i (0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N)\) is called the ith minor generalized eigenvalue, \((\cdot)^H\) stands for the conjugate transpose of a vector or a matrix, \(\delta_{i,j}\) is the Kronecker delta function. In this paper, \((R_y, R_z)\) and \(v_i\) are respectively called a matrix pencil and \(i\)th minor generalized eigenvector of \((R_y, R_z)\). The GHEP has been attracting great attention in many branches of signal processing, e.g., subspace tracking [1], [2], blind source separation [3], fault detection [4], pattern recognition [5], and array signal processing [6]–[10].

Adaptive estimation of \(v_i\) is also required and some adaptive estimators are proposed [11]–[13]. In our previous work [14], we proposed an adaptive estimator of the first minor (or principal) generalized eigenvector for fast and stable estimation. In [15], we also proposed the scheme for tracking the first \(r\) minor (or principal) generalized eigenvectors, which achieved the lower estimation error and the robustness against additive noise compared with [11]–[13] (see [15, Figs. 1 and 3]) while keeping the orthogonality. This scheme reduced estimation problem of \(v_i\) (\(i = 1, \ldots, r\)) of \((R_y, R_z)\) to that of the first minor generalized eigenvector \(v_1^{(i)} \in \mathbb{C}^{N-r+i}\) of a certain smaller matrix pencil \((R_y^{(i)}, R_z^{(i)})\) (\(i = 1, \ldots, r\)) by using the nested orthogonal complement structure (Section 2.1). The ith minor generalized eigenvector \(v_i\) can be calculated from \(v_1^{(i)}\) because \(v_1^{(i)}\) is an expression of \(v_1\) in a certain coordinate system \(\perp_i \in \mathbb{C}^{N \times (N-r+i)},\)

STABILIZATION OF ADAPTIVE EIGENVECTOR EXTRACTION
BY CONTINUATION IN NESTED ORTHOGONAL COMPLEMENT STRUCTURE

Kenji Kakimoto Daichi Kitahara Masao Yamagishi Isao Yamada

Department of Communications and Computer Engineering, Tokyo Institute of Technology, Japan

2. PRELIMINARIES
Let \(\mathbb{R}\) and \(\mathbb{C}\) be respectively the set of all real numbers and complex numbers. Capital and bold face small letters respectively express a matrix and a vector. The Euclidean norm and the \(B\)-norm of \(x := [x_1, x_2, \ldots, x_N]^T \in \mathbb{C}^N\) are respectively defined as \(\|x\| := \sqrt{\sum_{i=1}^{N}|x_i|^2}\) and \(\|x\|_B := \sqrt{x^H B x}, \) where \(B \in \mathbb{C}^{N \times N}\) is a Hermitian positive definite matrix and \((\cdot)^T\) stands for the transpose.

For self-containedness, we present a summary of the adaptive eigenvector extraction in [15].

2.1. Nested Orthogonal Complement Structure

Definition 1 (B-orthogonal complement matrix) Let \(B \in \mathbb{C}^{N \times N}\) be a Hermitian positive definite matrix. For \(u \in \mathbb{C}^N \setminus \{0\}\) \(U_\perp \in \mathbb{C}^{N \times (N-1)}\) is called a B-orthogonal complement matrix of \(u\) if
\[
U_\perp H B u = 0 \quad \text{and} \quad U_\perp H U_\perp = I_{N-1},
\]
where \(I_{N-1} \in \mathbb{R}^{(N-1) \times (N-1)}\) is the identity matrix.

Fact 1 (Example of a B-orthogonal complement matrix) One of B-orthogonal complement matrices of \(u\) can be calculated as
\[
U_\perp = \begin{bmatrix} I_{N-1} & -\frac{1}{\|u_\perp u_\perp^H\|} u_\perp u_\perp^H \\ \frac{1}{\|u_\perp u_\perp^H\|} u_\perp u_\perp^H \end{bmatrix},
\]
where \(u_\perp \in \mathbb{C}^{N-1}\) and \(u_\perp \in \mathbb{C}\) are respectively the first \((N-1)\) components and the last component of a normalized vector \(\tilde{u} := Bu/\|Bu\|,\) i.e., \(\tilde{u} := [u_\perp^T, u_\perp^H]^T\), and \(\theta : \mathbb{C} \rightarrow \mathbb{C}\) is defined as
\[
\theta(u_\perp) := \begin{cases} 1, & \text{if } u_\perp = 0; \\ \|u_\perp^H u_\perp\|, & \text{otherwise}. \end{cases}
\]
Algorithm 1 Normalized Quasi-Newton Algorithm [14]

With any $B$-normalized vector $u(0) \in C^N$ and $\lambda(0) \geq 0$, generate the sequence $(u(k), \lambda(k)) \in (C^N \setminus \{0\}) \times \mathbb{R}$ ($k = 0, 1, \ldots$) by

$$
\hat{u}(k + 1) := u(k) + \eta \left[ A^{-1}Bu(k)\lambda(k) + \nu^H(k)Au(k)(u(k)\lambda^{-1}(k) - 2u(k))\right]
$$

$$
u(k + 1) := \hat{u}(k + 1)/\|\hat{u}(k + 1)\|
$$

$$
\lambda(k + 1) := (1 - \gamma)\lambda(k) + \gamma u^H(k + 1)Au(k + 1)
$$

with step sizes $\eta > 0$ and $\gamma \in [0, 1]$, where $(u(k), \lambda(k))$ are the estimates of the first minor generalized eigenvector and eigenvalue.

Remark 1 The $B$-orthogonal complement matrix in (2) is not a trivial extension of the standard orthogonal complement matrix [16] satisfying $U^H_k U = 0$ and $U^H_k U_k = I_{N-k}$. In Definition 1, we use the condition $U^H_k U_k = I_{N-k}$ in place of a trivial one $U^H_k U_k = I_{N-k}$, which is a key for efficient computation of (3).

We reduce the estimation problem of the first minor generalized eigenvector $u_0(1) \in C^N$ of a matrix pencil $(A, B) \in C^{N \times N} \times C^{N \times N}$ to that of the first minor generalized eigenvector $u_0(1) \in C^{N-1}$ of a certain smaller matrix pencil $(A(1), B(1))$ as follows.

Fact 2 (Nested orthogonal complement structure) Define $N$ matrix pencils $(A(i), B(i))$ recursively as $(A(1), B(1)) := (A, B)$ and

$$
\begin{align*}
A(i+1) := & \left( U^H_1 A U_1(2) \right) U_2^{(i+1)} \in C^{(N-i) \times (N-i)}
B(i+1) := & \left( U^H_1 B U_1(2) \right) U_2^{(i+1)} \in C^{(N-i) \times (N-i)}
\end{align*}
$$

where $U_1^{(1)}$ is a $B(1)$-orthogonal complement matrix of the first minor generalized eigenvector $u_0^{(1)}$ of $(A(1), B(1))$. Then the $i$th minor generalized eigenvector $u_i^{(1)}$ of $(A, B)$ is expressed as

$$
u_i^{(1)} = U_1^{(i)} U_2^{(i+2)} \ldots U_2^{(i-1)} u_0^{(1)} = \perp_i u_0^{(1)}.
$$

In (5), we can regard $u^{(1)} \in C^{N-i-1}$ as an expression of $u^{(1)} \in C^N$ in the coordinate system $\perp_i \in C^{N \times (N-i)}$. By combining Fact 2 with Algorithm 1 (Normalized Quasi-Newton Algorithm [14]), which iteratively estimates the first minor generalized eigenvector, we establish the following scheme for estimation of the first minor generalized eigenvectors.

Scheme 1 (Extraction of the first minor generalized eigenvectors $u_0^{(1)}(i = 1, \ldots, r)$ of a matrix pencil $(A, B)$ by using Fact 2 [15])

1. Set $A(1) = A$ and $B(1) = B$.
2. For $i = 1, \ldots, r$
   a. Extract the first minor generalized eigenvector $u_0^{(1)}$ of the matrix pencil $(A(i), B(i))$.
   b. Compute a $B(i)$-orthogonal complement matrix $U_1^{(i)}$ of $u_0^{(1)}$.
   c. Set $A(i+1) := \left( U_1^{(i)} A U_2^{(i+1)} \right) U_2^{(i+1)} = \left( U_1^{(i)} B U_2^{(i+1)} \right) U_2^{(i+1)}$.
3. For $i = 2, \ldots, r$, compute $u_0^{(1)}$ by (5).

2.2. Adaptive Estimation of Generalized Eigenvectors

In many signal processing applications, the matrix pencil $(R_y, R_x)$ in (1) are defined as a pair of covariance matrices of input sequences $(y(k))_{k \geq 0}$ and $(x(k))_{k \geq 0}$, where $k$ denotes discrete time index. In adaptive case, we have to estimate the matrix pencil $(R_y, R_x)$ and its generalized eigenvectors simultaneously. As an adaptive version of Scheme 1, we proposed the following scheme [15] where $w_0^{(1)}(i)$ is the estimate of the $i$th minor generalized eigenvector $v_i$ at time $k$.

Scheme 2 (Adaptive version of Scheme 1)

1. Update the estimate $(R_y(1), R_x(1))$ of $(R_y, R_x)$.
2. For $i = 1, \ldots, r$
   a. Update the estimate $w_0^{(1)}(k)$, from $w_0^{(1)}(k-1)$, of the first minor generalized eigenvector of the matrix pencil $(R_y(1), R_x(1))$.
   b. Compute an $R_y(1)$-orthogonal complement matrix $W(1)(k)$ of $w_0^{(1)}(k)$.
   c. Set $R_y(i+1)(k) = (W(1)(k))^H R_y(1)(k) W(1)(k)$, $R_x(i+1)(k) = (W(1)(k))^H R_x(1)(k) W(1)(k)$.
3. For $i = 2, \ldots, r$, compute $w_0^{(1)}(k) = \left( \prod_{s=1}^{i-1} W(1)(k) \right) w_0^{(1)}(k)$.
4. $k \leftarrow k + 1$. Repeat 1–4 until $w_0^{(1)}(k) (i = 1, \ldots, r)$ converge.

In [15], $W(1)(k)$ in Scheme 2-2b is calculated as an adaptive version of (3), i.e.,

$$
W(1)(k) = \left[ I_{N - i} - \frac{\theta}{\lambda_{\text{up}}(k)} \left( \frac{\lambda_{\text{up}}(k)}{\lambda_{\text{low}}(k)} \right) \frac{\lambda_{\text{up}}(k)}{\lambda_{\text{low}}(k)} \right] - \theta \left( \frac{\lambda_{\text{up}}(k)}{\lambda_{\text{low}}(k)} \right) \frac{\lambda_{\text{up}}(k)}{\lambda_{\text{low}}(k)}
$$

where $\lambda_{\text{up}}(k) \in C^{N-i}$ and $\lambda_{\text{low}}(k) \in C$ satisfy $\lambda_{\text{up}}(k) = R_y(1)^{-1}(k) \lambda_{\text{up}}(k) / \|R_y(1)^{-1}(k) \lambda_{\text{up}}(k)\| = \left( \lambda_{\text{up}}(k)^2 - \lambda_{\text{low}}(k)^2 \right)$.

Moreover, an example of $(R_y(1), R_x(1))$ is shown in Section 4.

3. PROPOSED STABILIZATION OF ADAPTIVE EIGENVECTOR EXTRACTION

3.1. Instability in Time-Varying Coordinate System

We find that Scheme 2 has certain instability in the estimation of the non-first minor generalized eigenvectors $v_i (i = 2, \ldots, r)$. In Scheme 2, since the non-first minor generalized eigenvectors are estimated by using the time-varying coordinate systems

$$
\perp_i := \prod_{s=1}^{i-1} W(1)(k)
$$

as $w_0^{(1)}(k) = \perp_i w_0^{(1)}(k)$, the significant change of $\perp_i$ directly influences the estimation accuracy of $v_i$. To see this, we focus on Scheme 2-2a. In this step, if the time-varying coordinate systems change smoothly, i.e., $\perp_i(k-1) \approx \perp_i(k)$, then the estimation of $v_i$ based on the update of $w_0^{(1)}(k)$ from $w_0^{(1)}(k-1)$ is stable because $\perp_i(k-1) \approx \perp_i(k) w_0^{(1)}(k)$ can be expected. However, if the coordinate systems change significantly, i.e., $\perp_i(k-1) \not\approx \perp_i(k)$, then the estimation of $v_i$ based on the update of $w_0^{(1)}(k)$ from $w_0^{(1)}(k-1)$ is unstable because $\perp_i(k-1) w_0^{(1)}(k-1) \not\approx \perp_i(k) w_0^{(1)}(k)$.

Such a significant change of $\perp_i$ is based on those of $W(s)(k)$ ($s = 1, \ldots, i-1$) due to (7), and a significant change of $\perp(s)$ is caused by the discontinuity of $\theta$ in (4) (see Example 1 below).

Example 1 (Simple example of the significant change of $W(s)(k)$)

For simplicity, set $N = 2$, $R_y(1)(k) = R_x(1)(k) = I_2$, and $w_0^{(1)}(k) = \theta w_0^{(1)}(k) \in \mathbb{R}^2$ (in this case, $\theta w_0^{(1)}(k)$ returns the sign of $w_0^{(1)}(k)$). Figure 1 shows an example of the significant change of the orthogonal complement matrix $W(1)(k)$ (in this case, $W(1)(k) \in \mathbb{R}^{2 \times 1}$). In Fig. 1, green dotted line stands for $w_0^{(1)}(k) = (w_{\text{up}}(k), \delta)^T$ with a small positive value $\delta \in (0, \sqrt{2}/2]$, and red dotted line stands for $W(1)(k) = (\delta, -w_{\text{up}}(k))^T$ computed by (6). Suppose that $w_0^{(1)}(k)$ is updated.
Fig. 1. Example of significant change of $W^{(1)}_{\pm}(k) \in \mathbb{R}^{2 \times 1}$ in (6). to $w^{(1)}_{\pm}(k + 1) = (w_{u,p}(k), -\delta)^T$ (green solid line), and $W^{(1)}_{\pm}(k)$ is updated to $W^{(1)}_{\pm}(k + 1) = (\delta, w_{u,p}(k))^T$ (red solid line). Obviously, the orthogonal complement matrix changes significantly. In fact, the distance between $W^{(1)}_{\pm}(k + 1)$ and $W^{(1)}_{\pm}(k)$

$$\|W^{(1)}_{\pm}(k + 1) - W^{(1)}_{\pm}(k)\| = 2\sqrt{w^{2}_{\text{up}}(k)} = 2\sqrt{1 - \delta^2}$$

implies that $W^{(1)}_{\pm}(k)$ changes significantly as $\delta \to +0$ while $\|w^{(1)}_{\pm}(k + 1) - w^{(1)}_{\pm}(k)\| = 2\delta \to +0$ as $\delta \to +0$.

On the other hand, the absolute value of the inner product between $W^{(1)}_{\pm}(k + 1)$ and $W^{(1)}_{\pm}(k)$

$$|\langle W^{(1)}_{\pm}(k + 1), W^{(1)}_{\pm}(k) \rangle| = |\delta^2 - w^{2}_{\text{up}}(k)| = 1 - 2\delta^2.$$

approaches 1 as $\delta \to +0$, which means that the range spaces of $W^{(1)}_{\pm}(k + 1)$ and $W^{(1)}_{\pm}(k)$ approach the same subspace as $\delta \to +0$.

3.2. Continuation in Nested Orthogonal Complement Structure

We stabilize Scheme 2 by maximally utilizing the information of the previous estimate $w^{(1)}_{\pm}(k - 1) = \perp(k - 1)w^{(1)}_{\pm}(k - 1)$ for the update of $w^{(1)}_{\pm}(k)$. From Example 1, even if the orthogonal complement matrix $W^{(s)}_{\pm}(k)$ $(s = 1, \ldots, i - 1)$ changes significantly, its range space is expected to change smoothly. Therefore, the range space of the time-varying coordinate system $\perp(k)$ is also expected to change smoothly from (7). This motivates the use of the expression which stands for the best approximation of the previous estimate $w^{(1)}_{\pm}(k - 1)$ in the current coordinate system $\perp(k)$, i.e.,

$$\arg\min_{w \in \mathbb{R}^{\perp(k)}} \|w^{(1)}_{\pm}(k - 1) - w\|^2 \quad = \perp(k) \left(\perp(k)^H \perp(k)\right)^{-1} \perp(k)^H w^{(1)}_{\pm}(k - 1) \quad (8)$$

$$= \perp(k) \left(\perp(k)^H w^{(1)}_{\pm}(k - 1)\right), \quad (9)$$

where $\mathbb{R}^{\perp(k)}$ is the range space of $\perp(k)$, (8) is derived as the projection of $w^{(1)}_{\pm}(k - 1)$ onto $\mathbb{R}^{\perp(k)}$, and (9) is derived from (2) and (7). Since $w^{(1)}_{\pm}(k - 1) = \perp(k - 1)w^{(1)}_{\pm}(k - 1) \approx \perp(k) \perp(k)^H w^{(1)}_{\pm}(k - 1)$, it is stable to update $w^{(1)}_{\pm}(k)$ from $(\perp(k)^H w^{(1)}_{\pm}(k - 1))$ instead of $w^{(1)}_{\pm}(k - 1)$. More precisely, we propose to update $w^{(1)}_{\pm}(k)$ from an $R^{(1)}_{\pm}$-normalized vector

$$\tilde{w}^{(1)}_{\pm}(k) := \frac{\perp(k)^H w^{(1)}_{\pm}(k - 1)}{\|\perp(k)^H w^{(1)}_{\pm}(k - 1)\|^2} R^{(1)}_{\pm}(k)$$

instead of $w^{(1)}_{\pm}(k - 1)$, and we propose the following scheme.

**Scheme 3** (Proposed scheme for stable adaptive estimation)

1. Update the estimate $(R^{(1)}_{y}(k), R^{(1)}_{x}(k))$ of $(R_{y}, R_{x})$.
2. For $i = 1, \ldots, r$
   a. If $i \geq 2$, compute $\tilde{w}^{(i)}_{\pm}(k - 1)$ by (10).
   
   b. Update the estimate $w^{(i)}_{\pm}(k)$, from $w^{(i)}_{\pm}(k - 1)$, of the first minor generalized eigenvector of the matrix pencil $(R^{(i)}_{y}(k), R^{(i)}_{x}(k))$.
   c. Compute an $R^{(i)}_{y}$-orthogonal complement matrix $W^{(i)}_{\pm}(k)$ of $w^{(i)}_{\pm}(k)$.
   d. Set $R^{(i+1)}_{y}(k) = (W^{(i)}_{\pm}(k))^H R^{(i)}_{y}(k) W^{(i)}_{\pm}(k)$, $R^{(i+1)}_{x}(k) = (W^{(i)}_{\pm}(k))^H R^{(i)}_{x}(k) W^{(i)}_{\pm}(k)$.

3. For $i = 2, \ldots, r$, compute $w^{(i)}_{\pm}(k) = \prod_{j=1}^{i-1} W^{(j)}_{\pm}(k) w^{(i)}_{\pm}(k)$.

4. $k \leftarrow k + 1$. Repeat 1–4 until $w^{(i)}_{\pm}(k)$ converge.

3.3. Additional Complexity of Proposed Stabilization

The complexity of Scheme 3 does not increase so much compared with Scheme 2. To see this, we shall observe the complexity of Scheme 3-2a. The calculation of (10) can be separated into two steps, $(\perp(k))^H w^{(1)}_{\pm}(k - 1)$ and $R^{(1)}_{\pm}$-normalization. From (7),

$$\perp(k)^H w^{(1)}_{\pm}(k - 1) = (W^{(1)}_{\pm}(k))^H \perp(k)^H w^{(1)}_{\pm}(k - 1), \quad (11)$$

which requires $(i - 1)$ times multiplications of vectors and orthogonal complement matrices. Fortunately, we can calculate each multiplication with low complexity by using (6). In the following discussion, time index $k$ is omitted for simplicity. For $W^{(s)}_{\pm} \in \mathbb{C}^{(N - s)(N - s)}$ and any $t = t^{(s)}_{\text{up}}, t^{(s)}_{\text{low}} \in \mathbb{C}^{N - s}$, we have

$$W^{(s)}_{\pm} t = \left[ I_{N - s} - \frac{\tilde{w}^{(s)}_{\text{up}}(k)}{1 + \tilde{w}^{(s)}_{\text{low}}(k)} \frac{\tilde{w}^{(s)}_{\text{up}}(k)}{1 + \tilde{w}^{(s)}_{\text{low}}(k)} \right] t^{(s)}_{\text{up}} - \frac{\tilde{w}^{(s)}_{\text{up}}(k)}{1 + \tilde{w}^{(s)}_{\text{low}}(k)} t^{(s)}_{\text{low}} \frac{\tilde{w}^{(s)}_{\text{up}}(k)}{1 + \tilde{w}^{(s)}_{\text{low}}(k)} \tilde{w}^{(s)}_{\text{up}}(k). \quad (12)$$

Since (12) can be computed with $O(2(N - s))$ multiplications, by counting all multiplications over $s = 1, \ldots, i - 1$, the calculation of (11) requires $O(2N - i^2)$ multiplications.

The $R^{(1)}_{\pm}$-normalization needs $O((N - i + 1)^2)$ multiplications for calculation of $w^{(i)}_{\pm} R^{(i)}_{\pm} w^{(i)}_{\pm}$. Since (10) is calculated for $i = 2, \ldots, r$, the total computational complexity of Scheme 3-2a is $O(N^2 - r^2/6) + O(N^2 - N^2 + r^2/6) = O(N^2 r)$. This computational complexity is much smaller than the total computational complexity $O(12N^2 r - 11N^2 r^2 + 11r^3/6)$ of Scheme 2 [15].

4. APPLICATION TO SUBSPACE TRACKING

In Scheme 2 and Scheme 3, we estimate the matrix pencil $(R_y, R_x)$ as a pair of exponential weighted sample covariance matrices

$$\begin{bmatrix} R^{(1)}_{y}(k) = \beta R^{(1)}_{y}(k - 1) + y(k) y(k)^H \\ R^{(1)}_{x}(k) = \alpha R^{(1)}_{x}(k - 1) + x(k) x(k)^H \end{bmatrix}$$

with forgetting factors $\alpha, \beta \in (0, 1)$.

4.1. Performance Criteria

In a scenario of application to subspace tracking, we evaluate the performance of Scheme 3 (proposed scheme) compared with Scheme 2 through $L = 100$ independent runs. For comparison, we observe the similarity between $u_i$ and $w^{(i)}_{\pm}(k)$ (the $i$th minor generalized eigenvector of $(R_y, R_x)$ and the estimate of $u_i$ at time $k$ in the $i$th independent run) in terms of Direction Cosine and its average

$$\text{DC}_{i,j}(k) := \frac{\|w^{(i)}_{\pm}(k)\|^2}{\|w^{(i)}_{\pm}(k)\|^2} \text{ and } \text{ADDC}_{i,j}(k) := \frac{1}{r} \sum_{i=1}^{r} \text{DC}_{i,j}(k).$$
Define the averages of $\text{DC}_{i,j}(k)$ and $\text{ADC}_{i,j}(k)$ in $L$ independent runs as $\overline{\text{DC}}_{i,j}(k) := \frac{1}{L} \sum_{j=1}^{L} \text{DC}_{i,j}(k)$ and $\overline{\text{ADC}}(k) := \frac{1}{L} \sum_{j=1}^{L} \text{ADC}_{i,j}(k)$. We also measure numerical stabilities by two kinds of Sample Standard Deviations

$$\text{SSD}_i(k) := \sqrt{\frac{1}{L-1} \sum_{j=1}^{L} (\text{DC}_{i,j}(k) - \overline{\text{DC}}_{i,j}(k))^2}$$

and

$$\text{SSD}(k) := \sqrt{\frac{1}{L-1} \sum_{j=1}^{L} (\text{ADC}_{i,j}(k) - \overline{\text{ADC}}(k))^2}.$$

### 4.2. Numerical Experiments

The input samples are generated by

$$y(k) = \sqrt{2} \sin(0.37\pi k + \theta_1) + n_1(k)$$

and

$$x(k) = \sqrt{2} \sin(0.42\pi k + \theta_2) + \sqrt{2} \sin(0.65\pi k + \theta_3) + n_2(k),$$

where the initial phase $\theta_i$ ($i = 1, 2, 3$) has the uniform distribution in $[0, 2\pi]$, $n_1(k)$ and $n_2(k)$ are white Gaussian noise with variance $\sigma^2 = 0.1$. The input vectors $y(k) \in \mathbb{R}^N$ and $x(k) \in \mathbb{R}^N$ ($N = 8$) are defined as $y(k) := (y(k), y(k-1), \ldots, y(k-N+1))^T$ and $x(k) := (x(k), x(k-1), \ldots, x(k-N+1))^T$ ($k \geq N$).

We adaptively estimate the first four ($r = 4$) minor generalized eigenvectors $v_i$ ($i = 1, 2, 3, 4$) of the matrix pencil $(R_y, R_x)$ with

\begin{align*}
(R_y)_{i,j} &:= \cos(0.37\pi(j-i)) + \delta_{ij}\sigma^2, \\
(R_x)_{i,j} &:= \cos(0.42\pi(j-i)) + \cos(0.65\pi(j-i)) + \delta_{ij}\sigma^2,
\end{align*}

where $(\cdot)_{i,j}$ stands for the $(i,j)$-component of the matrix. These matrices are used for computing true generalized eigenvectors $v_i$, the parameters $\alpha = \beta = 0.998$, $\eta = 1/(\lambda_N/\lambda_1 - 1)$, $\gamma = 0.998$ and the initial estimates $R_y^{(1)}(0) = R_x^{(1)}(0) = I_N$.

Figure 2 shows one of the outcomes in $L (= 100)$ independent runs. Figure 2(a) depicts $\theta(\bar{w}^{(1)}_{low}(i))$ ($i = 1, 2, 3$). In this case, $\theta(\bar{w}^{(1)}_{low})$ returns the sign of $\bar{w}^{(1)}_{low}$. Figures 2(b) and 2(c) respectively depict $\text{DC}_{i,j}$ ($i = 1, 2, 3, 4$) of Scheme 2 and Scheme 3. From Fig. 2, we observe that when $\theta(\bar{w}^{(1)}_{low})$ changes significantly i.e., when the sign of $\bar{w}^{(1)}_{low}$ changes from positive/negative to negative/positive, the estimation by Scheme 2 is unstable, as mentioned in Section 3.1, while the estimation by Scheme 3 is stable. Figure 3 shows the result in 100 independent runs. Figures 3(a), 3(b) and 3(c) respectively depict $\overline{\text{ADC}}$, SSD$_1$, and SSD of Scheme 2 and Scheme 3 (note that in Fig. 3(b), there is no difference between Scheme 2 and Scheme 3 for the estimation of the first minor generalized eigenvector $v_i$). From these figures, we observe that Scheme 3 has better numerical stability than Scheme 2. Especially, SSD$_1$ ($i = 3, 4$) in Fig. 3(b) are significantly improved.

### 5. CONCLUSIONS

We found that the numerical instability of the original adaptive eigenvector extraction (Scheme 2) is caused by the significant change of the time-varying coordinate systems. To cope with the significant change, we proposed to use the expression of the nearest vector from the previous estimate in the current coordinate system. This proposed stabilization is realized by the projection of the previous estimate onto the range space of the current coordinate system, and its complexity is low compared with the total complexity of the original scheme. Numerical experiments showed the excellent performance of the proposed stabilization.

**Acknowledgement** We would like to thank Tuan Duong Nguyen (NTT Communications) for fruitful discussions.
6. REFERENCES


